Inner-Product Spaces

Let V be a vector space over $F = \mathbb{R}$ or $F = \mathbb{C}$, finite or infinite-dimensional. An inner product on V is a function $\langle \cdot, \cdot \rangle : V \times V \to F$ which satisfies the following axioms:

1. **Positive-Definiteness:** $\langle v, v \rangle \geq 0$ for all $v \in V$, and $\langle v, v \rangle = 0$ if and only if $v = 0_V$.

2. **Multilinearity:** $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$ and $\langle av, w \rangle = a\langle v, w \rangle$ for all $v, v', w \in V$ and $a \in W$.

3. **Conjugate Symmetry:** $\langle w, v \rangle = \overline{\langle v, w \rangle}$ for all $v, w \in V$.

Inner-Product Space: A real or complex vector space V equipped with an inner-product.

Note that axioms 2 and 3 imply $\langle v, aw \rangle = \overline{a}\langle v, w \rangle$ and $\langle v, w + w' \rangle = \langle v, w \rangle + \langle v, w' \rangle$ for all $v, w \in V$ and $a \in F$.

Examples: The usual dot product on $\mathbb{R}^n$, the generalized dot product on $\mathbb{C}^n$, the inner-product on $P([a, b])$ defined by $\langle f, g \rangle = \int_a^b f(x)g(x)\ dx$.

Norm: $||v|| = \sqrt{\langle v, v \rangle}$. This satisfies $||av|| = |a| \cdot ||v||$ where $|a| = \sqrt{a\overline{a}}$ is absolute value (if real) or length (if complex).

Orthogonal vectors: $u_1, \ldots, u_n$ are mutually orthogonal iff $\langle u_i, u_j \rangle = 0$ for all $i \neq j$.

Orthonormal vectors: $u_1, \ldots, u_n$ are mutually orthonormal iff $\langle u_i, u_j \rangle = \delta_{i,j}$ for all i, j. In other words, they are mutually orthogonal and have length 1.

Orthonormal projection: Let $u_1, \ldots, u_n$ be mutually orthonormal. Let $U = \text{span}(u_1, \ldots, u_n)$. The linear operator $P : V \to U$ defined by $Pv = \sum \langle v, u_i \rangle u_i$ is called orthonormal projection onto U.

Properties of orthogonal and orthonormal vectors:

1. Mutually orthogonal vectors $u_1, \ldots, u_n$ are linearly independent.

Proof: Suppose $\sum a_i u_i = 0_V$. Taking the inner product with u_j we obtain $0 = \langle 0_V, u_j \rangle = \langle \sum a_i u_i, u_j \rangle = \sum a_i \langle u_i, u_j \rangle = a_j ||u_j||^2 = a_j$.

2. Let $u_1, \ldots, u_n$ be mutually orthogonal. Then $|| \sum u_i ||^2 = \sum ||u_i||^2$. This is called the Pythagorean Theorem.

Proof: $\langle \sum u_i, \sum u_i \rangle = \sum \langle u_i, u_j \rangle = \sum \langle u_i, u_i \rangle$.

3. Let $u_1, \ldots, u_n$ be mutually orthonormal. Then $|| \sum a_i u_i || = \sqrt{\sum |a_i|^2}$.

Proof: $\langle \sum a_i u_i, \sum a_i u_i \rangle = \sum a_i \overline{a_j} \langle u_i, u_j \rangle = \sum a_i \overline{a_i}$.

4. Let $u_1, \ldots, u_n$ be mutually orthonormal. Let $U = \text{span}(u_1, \ldots, u_n)$. Then for any $u \in U$, $u = \sum \langle u, u_i \rangle u_i$. In other words, $u = Pu$ where P is orthonormal projection onto U. This also implies $P^2 = P$.

Proof: Write $u = \sum a_i u_i$. Then $\langle u, u_j \rangle = \langle \sum a_i u_i, u_i \rangle = \sum a_i \langle u_i, u_j \rangle = a_j$.

Properties of orthonormal projection:

1. Let $u_1, \ldots, u_n$ be mutually orthonormal. Let $U = \text{span}(u_1, \ldots, u_n)$. Then for any $v \in V$ and for any $u \in U$, $v - Pv$ and u are orthogonal to each other, where P is orthonormal projection onto U.

Proof: For any j, $\langle Pv, u_j \rangle = \langle \sum \langle v, u_i \rangle u_i, u_j \rangle = \sum \langle v, u_i \rangle \langle u_i, u_j \rangle = \langle v, u_j \rangle$. Subtracting, $\langle v - Pv, u_j \rangle = 0$.

2. Let $u_1, \ldots, u_n$ be mutually orthonormal. Let $U = \text{span}(u_1, \ldots, u_n)$. Then for any $v \in V$, the unique vector $u \in U$ that minimizes $||v - u||$ is Pv.

Proof: Let $u \in U$ be given. Then we know that $v - Pv$ and $Pv - u$ are orthogonal to each other. By the Pythagorean Theorem, $||v - u||^2 = ||v - Pv||^2 + ||Pv - u||^2 \geq ||v - Pv||^2$, with equality iff $||Pv - u|| = 0$ iff $u = Pv$.

Theorem: Every finite-dimensional subspace of an inner product space has an orthonormal basis.

Proof: Let V be the inner product space. Let U be a subspace of dimension n. We prove that U has an orthonormal basis by induction on n.

Base Case: $n = 1$. Let $\{u_1\}$ be a basis for U. Then $\{\frac{u_1}{||u_1||}\}$ is an orthonormal basis for U.

Induction Hypothesis: If U has dimension n then it has an orthonormal basis $\{u_1, \ldots, u_n\}$.

Inductive Step: Let U be a subspace of dimension $n+1$. Let $\{v_1, \ldots, v_{n+1}\}$ be a basis for U. Write $U_n = \text{span}(v_1, \ldots, v_n)$. By the induction hypothesis, U_n has an orthonormal basis $\{u_1, \ldots, u_n\}$. Let P be orthonormal projection onto U_n. Then the vectors $u_1, \ldots, u_n, v_{n+1} - Pv_{n+1}$ are mutually orthogonal and form a basis for U. Setting

$$u_{n+1} = \frac{v_{n+1} - Pv_{n+1}}{||v_{n+1} - Pv_{n+1}||},$$

the vectors $u_1, \ldots, u_{n+1}$ form an orthonormal basis for U.

Remark: The proof of this last theorem provides an algorithm (Gram-Schmidt) for producing an orthonormal basis for a finite-dimensional subspace U: Start with any basis $\{v_1, \ldots, v_n\}$. Set $u_1 = \frac{v_1}{||v_1||}$. This is an orthonormal basis for $\text{span}(v_1)$. Having found an orthonormal basis $\{u_1, \ldots, u_k\}$ for $\text{span}(v_1, \ldots, v_k)$, one can produce an orthonormal basis for $\text{span}(v_1, \ldots, v_{k+1})$ by appending the vector

$$u_{k+1} = \frac{v_{k+1} - Pv_{k+1}}{||v_{k+1} - Pv_{k+1}||},$$

where P is orthonormal projection onto $u_1, \ldots, u_k$.

A Minimization Problem: Consider the problem of finding the best polynomial approximation $p(x) \in P_5([-\pi, \pi])$ of $\sin x$, where by best we mean that

$$\int_{-\pi}^{\pi} (\sin x - p(x))^2 \; dx$$

is a small as possible. To place this in an inner-product setting, we consider $P_5([-\pi, \pi])$ to be a subspace of $C([-\pi, \pi])$, where the latter is the vector space of continuous functions from $[-\pi, \pi]$ to $\mathbb{R}$. Then $C([-\pi, \pi])$ has inner product defined by $\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x) \; dx$. We are trying to minimize $||\sin x - p(x)||^2$. However, we know how to minimize $||\sin x - p(x)||$: $p(x) = P(\sin x)$ where P is orthogonal projection onto the finite-dimensional subspace $P_5([-\pi, \pi])$. The latter has basis

$$\{1, x, x^2, x^3, x^4, x^5\},$$

and Gram-Schmidt can be applied to produce an orthonormal basis

$$\{u_0(x), u_1(x), u_2(x), u_3(x), u_4(x), u_5(x)\}.$$

Therefore the best polynomial approximation is $\sum \alpha_i u_i(x)$ where

$$\alpha_i = \langle \sin x, u_i(x) \rangle = \int_{-\pi}^{\pi} \sin x \cdot u_i(x) \; dx.$$

The approximation to $\sin x$ given in the book on page 115 is

$$\frac{x}{1.01229} - \frac{x^3}{6.44035} + \frac{x^3}{177.207},$$

in contrast to the Taylor Polynomial

$$\frac{x}{1} - \frac{x^3}{6} + \frac{x^5}{120}.$$

Cauchy-Schwarz Inequality: $|\langle u, v\rangle| \leq ||u|| \cdot ||v||$.

Proof: We have equality when u and v are linearly dependent. Now suppose u and v are linearly independent. Let $\{e_1, e_2\}$ be an orthonormal basis for span(u, v) and write $u = a_1e_1 + a_2e_2$ and $v = b_1e_1 + b_2e_2$. Then $|\langle u, v\rangle| = |a_1\overline{b_1} + a_2\overline{b_2}|$ and $||u|| \cdot ||v|| = \sqrt{a_1\overline{a_1} + a_2\overline{a_2}}\sqrt{b_1\overline{b_1} + b_2\overline{b_2}}$, and we are reduced to proving the complex inequality

$$|a_1\overline{b_1} + a_2\overline{b_2}| \leq \sqrt{a_1\overline{a_1} + a_2\overline{a_2}}\sqrt{b_1\overline{b_1} + b_2\overline{b_2}}.$$

Squaring both sides, this is equivalent to

$$(a_1\overline{b_1} + a_2\overline{b_2})(\overline{a_1}b_1 + \overline{a_2}b_2) \leq (a_1\overline{a_1} + a_2\overline{a_2})(b_1\overline{b_1} + b_2\overline{b_2}),$$

which is equivalent to

$$a_1\overline{b_1}\overline{a_2}b_2 + a_2\overline{b_2}\overline{a_1}b_1 \leq a_1\overline{a_1}b_2\overline{b_2} + a_2\overline{a_2}b_1\overline{b_1},$$

which is equivalent to

$$a_1\overline{a_1}b_2\overline{b_2} + a_2\overline{a_2}b_1\overline{b_1} - a_1\overline{b_1}\overline{a_2}b_2 - a_2\overline{b_2}\overline{a_1}b_1 \geq 0,$$

which is equivalent to

$$(a_1b_2 - a_2b_1)(\overline{a_1}\overline{b_2} - \overline{a_2}\overline{b_1}) \geq 0,$$

which is equivalent to

$$|a_1b_2 - a_2b_1|^2 \geq 0,$$

which is true.

Triangle Inequality: $||u+v|| \leq ||u|| + ||v||$.

Proof: Square both sides and subtract the left-hand side from the right-hand side. The result is

$$2||u||\cdot||v|| - \langle u, v\rangle - \langle v, u\rangle = 2||u||\cdot||v|| - 2\text{Re}\ \langle u, v\rangle \geq 2||u||\cdot||v|| - 2|\langle u, v\rangle| \geq 0$$

by Cauchy-Schwarz.

The Orthogonal Complement of a Subspace: Let V be a finite-dimensional inner-product space and let U be a subspace. We define

$$U^{\perp} = \{v \in V : \langle v, u\rangle = 0 \text{ for all } u \in U\}.$$

We can construct $U^{\perp}$ explicitly as follows: Let $\{u_1, \ldots, u_k\}$ be an orthonormal basis for U. Expand to an orthonormal basis $\{u_1, \ldots, u_n\}$ for V using Gram-Schmidt. The vectors in $\text{span}(u_{k+1}, \ldots, u_n)$ are orthogonal to the vectors in U. Moreover, for any $v \in U^{\perp}$, the coefficients of v in terms of the orthonormal basis are the inner product of v with each basis vector, which places $v \in \text{span}(u_{k+1}, \ldots, u_n)$. Therefore $U^{\perp} = \text{span}(u_{k+1}, \ldots, u_n)$. This immediately implies that $(U^{\perp})^{\perp} = \text{span}(u_1, \ldots, u_k) = U$. Note also that $V = U \bigoplus U^{\perp}$. To decompose a vector in V into something in U plus something in $U^{\perp}$ we can use $v = Pv + (v - Pv)$.

Linear Functionals and Adjoints: If V is a finite-dimensional inner-product space with orthonormal basis $\{e_1, \ldots, e_n\}$ and if $\phi : V \to F$ is a linear map then a simple calculation shows that $\phi(v)$ can be realized as inner-product with a fixed vector $v_0 = \sum_i \overline{\phi(e_i)} e_i$:

$$\phi(v) = \phi(\sum \langle v, e_i\rangle e_i) = \sum \langle v, e_i\rangle \phi(e_i) = \sum \langle v, \overline{\phi(e_i)} e_i\rangle = \langle v, v_0\rangle.$$

In particular, if W is a finite-dimensional inner-product space and $T : V \to W$ is a linear map and $w \in W$, then the linear functional $\phi_w : V \to F$ defined by $\phi_w(v) = \langle Tv, w\rangle$ satisfies

$$\phi_w(v) = \sum \langle v, \overline{\phi_w(e_i)} e_i\rangle = \langle v, \sum_i \overline{\langle Te_i, w\rangle} e_i\rangle$$

$$= \langle v, \sum_i \langle w, Te_i\rangle e_i\rangle.$$

This defines a linear map $T^* : W \to V$ via $T^*(w) = \sum_i \langle w, Te_i\rangle e_i$. In other words,

$$\langle Tv, w\rangle = \langle v, T^* w\rangle.$$

This gives rise to the properties listed on pp. 119–120.

The matrix representations of T and T^* with respect to orthonormal bases $\{e_i\}$ and $\{f_i\}$ are conjugate transposes of each other: Assume that $M(T) = (t_{ij})$ and $M(T^*) = (t^*_{ij})$. Then

$$\langle Te_i, f_j\rangle = \langle e_i, T^* f_j\rangle = \overline{\langle T^* f_j, e_i\rangle}$$

implies

$$t_{ij} = \overline{t^*_{ji}}$$

implies

$$t^*_{ij} = \overline{t_{ji}}.$$